

VẤN ĐỀ TỐI ƯU TRONG PHÂN HOẠCH NHẬN DẠNG ẢNH

THE OPTIMIZATION OF THE PICTORIAL RECOGNIZING SEGMENTATION

MAI HỘ
Trường Đại học Bách khoa, Đại học Đà Nẵng

TÓM TẮT

Bài báo giới thiệu phương pháp xây dựng hàm sai số trong phân hoạch nhận dạng ảnh với việc ứng dụng kiểu xác suất có điều kiện của Bayes. Do đó, nó thì dễ dàng xác định được giá trị nhỏ nhất của sai số và dẫn tới ảnh xử lý đảm bảo chất lượng theo yêu cầu của mọi lĩnh vực sử dụng.

ABSTRACT

A methode is necessary to creating of an error-function of the pictorial recognizing segmentation by applying the Bayes conditional probability. Thus, it is easy to determine the minimum error-value which leads to the concordent images processing meeting the demands of use.

1. Tổng quan về nhận dạng ảnh [1], [2], [3], [4]

Nhận dạng ảnh là một quá trình phân hoạch ảnh thành các đối tượng ảnh con, chúng được gán vào từng lớp nhãn để được đối sánh với mẫu và đối sánh theo các quy luật biết trước nào đó. Bấy giờ, có hai khái niệm được nói tới:

◇ Mỗi ảnh xử lý X là tập các ảnh con X_i ; biểu thức quan hệ của chúng được gọi là *không gian biểu diễn* các đối tượng của ảnh:

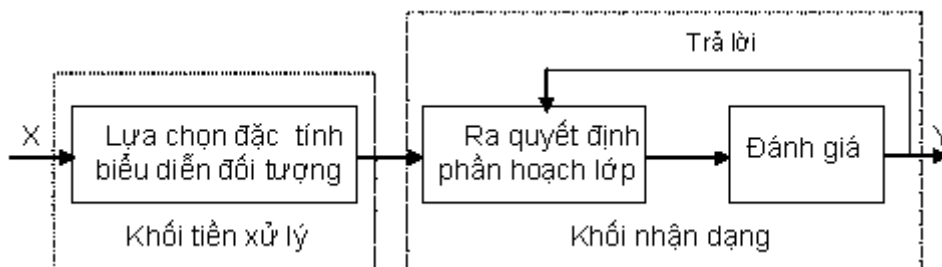
$$X = \{X_1, X_2, \dots, X_i, \dots, X_n\} \quad (\text{với } i=1..n) \quad (1)$$

Trong đó, mỗi ảnh con X_i chứa đựng một tập các vectơ khía cạnh x_{ii} :

$$X_i = \{x_{i1}, x_{i2}, \dots, x_{ii}, \dots, x_{in}\} \quad (2)$$

◇ Mỗi đối tượng ảnh con X_i có một tên nhãn w_i ; khi đó, tập các tên nhãn của các đối tượng được gọi là *không gian diễn dịch* của ảnh:

$$\Omega = \{w_1, w_2, \dots, w_i, \dots, w_n\} \quad (3)$$



Hình 1. Mô hình tổng quát hệ nhận dạng ảnh

Quá trình nhận dạng một đối tượng X là quá trình thực hiện phép ánh xạ F từ $X \rightarrow \Omega$, với F là tập các quy luật để xác định một phần tử trong X ứng với một định danh của nó ở trong Ω . Khi tập các quy luật F và tập tên định danh các đối tượng Ω biết trước, quá trình này được gọi là quá trình nhận dạng có giám sát hay còn gọi là học có giám sát (supervised learning); trường hợp ngược lại gọi là học không có giám sát (non supervised learning).

Nói chung, dù sử dụng kỹ thuật nào để nhận dạng ảnh, sơ đồ tổng quát của một hệ thống nhận dạng ảnh được chỉ ra trong hình 1 ở trên; trong đó, X là ảnh vào, Y là ảnh ra.

2. Đặc tả đối tượng trong phân hoạch ảnh [5], [6]

Mô hình nhận dạng là một kiểu đặc tả đối tượng trong quá trình phân tích và phân đoạn nhận dạng ảnh. Có hai kiểu mô hình đặc tả được sử dụng; đó là đặc tả theo tham số và đặc tả theo cấu trúc. Do đó, tương ứng với hai kiểu đặc tả này có hai kiểu phân hoạch nhận dạng sau đây.

Phân hoạch nhận dạng theo tham số:

Nội dung của kiểu mô hình này là sử dụng một vectơ để đặc tả đối tượng (ảnh con). Cụ thể, mỗi phần tử của vectơ mô tả được biểu thị một đặc tính của đối tượng; do đó, mỗi đặc tính này được biểu diễn bởi một hàm trục giao. Như vậy, ảnh nhận dạng được biểu diễn bởi một xâu các hàm trục giao. Tức là, giả sử ảnh con có đường biên bao C với $C(i)$ là một điểm bất kỳ trên đường biên này (với $i=1..N$) có tọa độ (x_i, y_i) , khi đó ta có:

✦ Tọa độ của trọng tâm ảnh con (x_0, y_0) được xác định:

$$x_0 = \frac{1}{N} \sum_{i=1}^N x_i \quad \text{và} \quad y_0 = \frac{1}{N} \sum_{i=1}^N y_i \quad (4)$$

✦ Môment trung tâm bậc N (m_N) của ảnh con bao bởi đường biên C được tính:

$$m_N = \frac{1}{N} \sum_{i=1}^N [(x_i - x_0)(y_i - y_0)]^N \quad (5)$$

Phân hoạch nhận dạng theo cấu trúc:

Cách tiếp cận theo kiểu này được sử dụng để mô tả đối tượng nhận dạng bởi một số khái niệm biểu thị đối tượng cơ sở trong ngôn ngữ tự nhiên. Điều đó được hiểu, một đối tượng nhận dạng bao giờ cũng có cấu trúc là một tập các đối tượng đơn giản (điểm, đoạn thẳng, đoạn cong...). Một cách hình thức, người ta coi mô hình này tương đương với một văn phạm G có cấu trúc:

$$G = (V_T, V_N, P, S) \quad (6)$$

Ở đây, G là hàm mô tả cấu trúc đối tượng, V_T là bộ ký hiệu kết thúc, V_N là bộ ký hiệu không kết thúc, P là luật sản sinh và S biểu thị dạng ký hiệu bắt đầu.

3. Bản chất của nhận dạng đối tượng ảnh [5]

Bản chất của nhận dạng đối tượng ảnh là một quá trình đối sánh ảnh gồm ba giai đoạn: 1. Lựa chọn mô hình biểu diễn đối tượng; 2. Lựa chọn phương pháp nhận dạng (còn gọi: lựa chọn phương pháp ra quyết định); 3. Suy diễn quá trình học, còn gọi là quá trình luyện học nhận dạng.

Nói một cách dễ hiểu, bản chất của việc đối sánh ảnh chính là tìm ra quy luật và thuật toán để gán cho đối tượng một cái tên xác định. Trong đó, công việc suy diễn quá trình luyện học là quan trọng. Quá trình này có hai dạng, đó là học có giám sát và học không có giám sát.

Luyện học có giám sát là phương pháp nhận dạng dựa trên các kiến thức biết trước để thực hiện thành công các thao tác phân biệt và phân hoạch ảnh. Đặc điểm mấu chốt của kỹ thuật này là sử dụng một thư viện chứa đựng các mẫu chuẩn có sẵn. Đối tượng ảnh cần nhận dạng được đối sánh với mẫu chuẩn để xem nó thuộc loại nào. Khi đó công việc chủ yếu là thiết kế một hệ thống để có thể đối sánh đối tượng ảnh với mẫu chuẩn và ra quyết định gán chúng vào một lớp nhãn tương ứng.

Luyện học không có giám sát là kỹ thuật nhận dạng phải tự định ra các lớp khác nhau và phải xác định các tham số đặc trưng cho từng lớp. Theo phương pháp học này, thì, số lớp và các đặc trưng của từng lớp không được biết trước; cho nên, công việc đối sánh ảnh phải tiến hành bằng mọi cách nhằm gộp nhóm có thể và chọn lựa cách phân tích và phân đoạn tốt nhất. Nghĩa là: bắt đầu từ tập dữ liệu, phải thiết kế nhiều thủ tục xử lý khác nhau nhằm phân lớp và nâng cấp dần để đạt được một phương án xử lý nhận dạng hiệu quả.

4. Vấn đề tối ưu trong phân hoạch nhận dạng không gian biểu diễn ảnh [5]

4.1. Khái quát về phân hoạch không gian

◇ Phân hoạch không gian:

Giả sử ảnh xử lý có không gian biểu diễn đối tượng của ảnh là $X = \{X_i \text{ với } i=1..N\}$. Người ta nói: P là một phân hoạch của không gian X thành các lớp C_i và C_j (i và j là chỉ số của vùng phân hoạch) thuộc X, nếu thỏa mãn:

$$(C_i \text{ AND } C_j) \text{ OR } C_i = X \quad (7)$$

Trường hợp lý tưởng, tập X được phân hoạch hoàn toàn. Song, trong thực tế không gian biểu diễn đối tượng chỉ phân hoạch được từng phần. Do đó, việc phân tích và phân đoạn nhận dạng phải dựa vào việc xây dựng một ánh xạ F từ $X \rightarrow P$. Công cụ để xây dựng ánh xạ này là các hàm ra quyết định phân lớp.

◇ Hàm ra quyết định phân lớp:

Để phân hoạch các đối tượng vào các lớp, người ta phải xác định số lớp và ranh giới giữa các lớp đó; do đó hàm phân hoạch còn được gọi là hàm phân lớp hay hàm phân biệt các lớp. Gọi $\{g_i\}$ là tập các hàm phân hoạch lớp, nó được định nghĩa:

Nếu $g_k(X) > g_i(X)$ với $\forall i \neq k$ thì quyết định X thuộc lớp k ($X \in \text{lớp } k$). Điều đó cho thấy, để phân biệt k lớp thì cần phải có (k-1) hàm phân biệt. Trong xử lý ảnh, hàm phân biệt g của một lớp nào đó được sử dụng là một hàm tuyến tính có dạng:

$$g(X) = w_0 + w_1 X_1 + w_2 X_2 + \dots + w_k X_k \quad (8)$$

Ở đây: w_i là trọng số gán cho các thành phần X_i ; w_0 là trọng số viết gọn ban đầu.

4.2. Ứng dụng phương pháp Bayes để xác định sai số phân hoạch ảnh

Phương pháp Bayes được ứng dụng để thực hiện bài toán phân hoạch có hiệu quả. Nếu các đối tượng nhận dạng tuân theo luật phân bố Gauss, với hàm mật độ xác suất cho bởi $P(X)$ là hàm mũ của cơ số tự nhiên:

$$P(X) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{(X-m)^2}{2\pi\sigma^2}\right) \quad (9)$$

(Ở đây: $\pi=3.14$; tập ảnh X, moment trung tâm m và bán kính gauss σ).

Khi đó, phương pháp Bayes dựa vào xác suất có điều kiện để xây dựng hàm phân hoạch cho đối tượng: Gọi $P(X/C_i)$ là xác suất để có X xuất hiện tại lớp C_i ; $P(C_i/X)$ là xác suất có điều kiện để X thuộc lớp C_i (với X là đối tượng nhận dạng và C_i là các lớp đối tượng). Công thức tính xác suất có điều kiện của Bayes được viết:

$$P(C_i/X) = \frac{P(X/C_i)P(C_i)}{P(X)} \quad (10)$$

Bây giờ phương pháp Bayes được diễn giải: Nếu ảnh nhận dạng cho trước các yếu tố:

- Không gian đối tượng: $X = \{X_i, i=1..N\}$ với $X_i = \{x_1, x_2, \dots, x_p\}$
- Không gian diễn dịch: $\Omega = \{C_1, C_2, \dots, C_r\}$ với C_i là lớp phân bố tế bào ảnh thứ i ($i=1..r$).

Quy tắc nhận dạng ảnh theo xác suất có điều kiện của Bayes được phát biểu:

Nếu ε là sai số của phép ánh xạ F từ X tới Ω sao cho $X \in C_k$ khi và chỉ khi $P(C_k/X) > P(C_i/X)$ $\forall i \neq k$ và $i=1..r$.

Trong thực tế, luôn luôn tồn tại sai số ε khi phân tích và phân đoạn nhận dạng ảnh. Do đó vấn đề tối ưu dữ liệu ảnh ở bước này là phải xây dựng quy tắc nhận dạng với sai số ε là nhỏ nhất.

4.3. Đánh giá và ra quyết định phân hoạch tối ưu

Muốn có được quyết định phân hoạch tối ưu, người ta phải xác định sai số ε của phép ánh xạ F (tức $F=(X \rightarrow \Omega)$) phải đạt giá trị tối thiểu.

Để xác định ε tối thiểu, phương pháp tiến hành: Muốn nhận được X thuộc lớp C_j ($X \in C_j$), ta phải xác định xác suất $P(C_j/X)$. khi đó sai số ε được tính:

$$\varepsilon = 1 - P(C_j/X) \quad (11)$$

Để tiện lợi, thường người ta dùng phương pháp đánh giá sai số trung bình. Cụ thể, người ta xây dựng một ma trận hai chiều $L[k,j]$ (với k và j là chỉ số của các lớp phân đoạn) có dấu được định nghĩa:

- Khi $L[k,j] > 0$ nếu $k \neq j$ thì có tồn tại sai số;
- Khi $L[k,j] \leq 0$ nếu $k = j$ thì không tồn tại sai số.

Nếu gọi $r_i(X)$ là sai số trung bình của sự phân hoạch theo lớp, nó được xác định:

$$r_j(X) = \sum_{j=1}^N L[k,j] P(C_j/X) \quad (12)$$

Ta có nhận xét: Để sai số của phép ánh xạ phân đoạn đạt giá trị ε_{\min} thì ta cần phải có r_{\min} . Từ các biểu thức (10) và (12), ta có:

$$r_j(X) = \sum_{j=1}^N L[k,j] P(X/C_j) P(C_j) \quad (13)$$

Bây giờ, ta xem xét quy trình tìm r_{\min} , được tiến hành như sau:

Giả sử hàm r (có đối số X) xác định liên tục trên khoảng $(a; b)$, có điểm $X_0 \in (a; b)$. Gọi $X_0 + \delta$ và $X_0 - \delta$ là các điểm lân cận của X_0 , với δ là khoảng lân cận phải và trái của X_0 , ký hiệu $V(\delta)$. Điểm X_0 được gọi là điểm cực tiểu của $r(X)$ nếu với mọi X thuộc vùng lân cận $V(\delta)$ thỏa mãn điều kiện:

$$r(X) > r(X_0) \text{ với } X \neq X_0 \quad (14)$$

Theo định lý về cực trị của hàm số, để tìm r_{\min} ta giải phương trình:

$$\frac{\partial r}{\partial X} \text{ xác định } X_0 \rightarrow \text{ tìm } r_{\min} = r_k(X_0) \quad (15)$$

Sau đó, theo phương pháp Bayes, quy tắc ra quyết định phân hoạch có tính đến giá trị min của sai số phép ánh xạ, cho phép đưa tới kết luận về bài toán phân hoạch ảnh tối ưu:

$$\text{Nếu } r_{\min} = r_k < r_j \text{ với } k < j \text{ khi } j=1..N \text{ thì } X_0 \in C_k \quad (16)$$

5. Kết luận

Đầu tiên hàm phân hoạch ảnh $r=f(X)$ được xác định bằng thực nghiệm với các bậc (số nguyên) từ thấp đến cao. Tiếp theo sử dụng các biểu thức (14), (15) và (16) để tiến hành phân tích và phân hoạch nhận dạng ảnh tối ưu. Từ những kết quả nghiên cứu đó, việc thiết kế giải thuật và xây dựng chương trình để thực thi các thao tác xử lý ảnh bằng phương pháp tối ưu phân hoạch nhận dạng ảnh này cho phép đạt được độ chính xác cao trong xử lý ảnh và góp phần thuận lợi cho quá trình xây dựng giao diện xử lý.

TÀI LIỆU THAM KHẢO

- [1] Dietrich Reschke, *Telematik*, Fachbuchverlage Leipzig, 2004.
- [2] A.K. Jain, *Fundamentals of Digital Image Processing*, Prentice Hall, 1996.
- [3] J. Pitas, *Digital Image Processing Algorithms*, Prentice Hall, 1998.
- [4] Mai Hộ, *Cơ sở Xử lý ảnh* (Giáo trình lưu hành nội bộ), Đà Nẵng, 2002.
- [5] Mai Hộ, *NC tối ưu dữ liệu tệp tin ảnh video*, đề tài KH&CN cấp Bộ, NT 2004.
- [6] Mai Hộ, *Xử lý ảnh bằng phép biến đổi cosin*, ĐHBK-ĐN, 2005, (Trang 205 Tuyển tập báo cáo HNKH, Kỷ niệm 30 năm thành lập trường).